# Bacteriophage Tree of Life

**Franklin L. Nobrega**, **Graça Pinto**, Joana Azeredo, Leon D. Kluskens

IBB – Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal

## Abstract

Viruses, in particular (bacterio)phages, are the most abundant entities on the planet with an estimate of 1031 phage particles. However, their official taxonomy is still based on morphotyping, mainly caused by the absence of genetic, and proteomic data [1].

The International Committee on Taxonomy of Viruses (ICTV) is assigned to regularly update taxonomic classification but about one-third of the classification of all tailed phages (96% of all known phages) remains restricted to the family. Even higher is the proportion of unclassified phages present in all phage groups [2]. Moreover, evolutionary relationships between phages with different particle morphology are not well studied. A perfect explanation for this fact is that phages are the main responsible for enhancing DNA transfer among prokaryotes with an estimated transduction of 1028 bp every year in each biome [3,4,5].

A reconstruction of evolutionary history becomes a difficult problem primarily due to the fast sequence drift and lack of omnipresent genes in phage genomes. Additionally, losses and recombinational exchanges of genes are so pervasive in phages that the plausibility of phylogenetic inference in phage kingdom has been questioned. With hundreds of completely sequenced phage genomes available, we undertook a systematic inventory of molecular characters providing a refinement in phage taxonomy by the identification of conserved biomolecular markers for each Family + Subfamily + Type, never found before [6]. In this work, we propose an evolutionary hypothesis using a matrix that includes biomolecular markers as vectors to infer the evolutionary history of phages and propose two molecular models for the mosaic gene evolution in phages (using MtArt+g and BLOSUM62+i as protein evolutionary algorithms). We compiled the profiles of presence and absence of 31 orthologous genes in 153 completely sequenced phages, representing all known ICTV classified phages to date, to construct the Bacteriophage Tree of Life. Conflicts between this hypothesis and other theories postulated are discussed.

## References

1. Calendar, R., The Bacteriophages, ed. R. Calendar. Vol. 2º. 2006: Oxford University Press;
2. Rohwer, F. and R. Edwards, The Phage Proteomic Tree: a Genome-Based Taxonomy for Phage. J Bacteriol, 2002;
3. Lima-Mendez, G., et al., Reticulate representation of evolutionary and functional relationships between phage genomes. Mol Biol Evol, 2008. 25(4): p. 762-77;
4. Glazko, G., et al., Evolutionary history of bacteriophages with double-stranded DNA genomes. Biol Direct, 2007. 2: p. 36;
5. Xu, J., R.W. Hendrix, and R.L. Duda, Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. Mol Cell, 2004. 16(1): p. 11-21;

6. Nóbrega, F. L. and Pinto, G. et al., A new look at bacteriophage classification. In preparation